5G 时代短视频新闻客户端如何构建智能视频能力

李 磊
（浙江在线技术中心，浙江 杭州 310039）

摘　要：随着 5G 时代的到来，短视频成为移动互联网的“风口”，牢牢抓住了用户的碎片化时间。短视频正逐步成为新闻客户端内容的重要表现形式，本文通过“短视频 +”人工智能技术，探索如何构建智能视频能力，赋能媒体深度融合。

关键词：5G；人工智能；短视频；新闻客户端；媒体融合　　中图分类号：G212　　文献标识码：A
文章编号：1671-0134（2021）02-036-04　　DOI：10.19483/j.cnki.11-4653/n.2021.02.007
本文著录格式：李磊 .5G 时代短视频新闻客户端如何构建智能视频能力 [J]. 中国传媒科技，2021（02）：36-38+110.

1. 短视频新闻客户端现状

近年来，抖音、快手、微视等短视频社交平台发展迅猛，同时资讯类短视频也快速兴起。QuestMobile 发布的《中国移动互联网 2020 年半年大报告》中显示，短视频行业月活跃用户规模在 2020 年 6 月已达到 8.52 亿。[1] 随着音视频技术的发展和 5G 网络的大面积覆盖，1 分钟以内的短视频内容牢牢地抓住了高速移动互联网时代下用户的碎片化时间，并且让用户长时间沉迷于短视频的观看，成为现象级“霸屏”产品。其内容形式特点在于，有鲜明的人设，短小精悍且有趣，能够与用户产生情感共鸣的，从而直接触发用户互动性强，黏性高。

通过短视频来呈现新闻报道相较于传统的图文、广电视频新闻等的报道形式，短视频新闻创新在于将视、听觉融为一体，视频进入“秒”时代，以其碎片化、移动化、社交化的特征，迎合了移动互联网时代的用户碎片化、浅阅读、强交互的内容消费习惯，同时将最精彩、最能表达情感、最能引起共鸣的内容直接触达给用户。随着媒体融合向纵深发展，围绕打造“四全媒体”的要求，短视频作为资讯内容模式逐步成为传播的重要手段，短视频新闻为传统媒体实现深度融合的关键切入点，为媒体融合升级转型提供了全新的驱动力。鉴于短视频在内容传播上的趋势，各大媒体纷纷推出了短视频新闻业务和频道，2016 年 10 月新京报上线了“我们”视频项目，同年 11 月梨视频客户端的上线，形成了早期的资讯类短视频平台的雏形。近两年，随着进一步推动媒体深度融合，央媒和地方媒体投入重要资源打造短视频客户端，如人民日报社推出主流短视频 PUGC 聚合平台“人民日报 +”客户端，中央广播电视总台推出的综合性视听新媒体旗舰平台“央视频”客户端，浙江在线倾力打造的服务长三角一体化发展国家战略的短视频新闻客户端“天目新闻”，标志着短视频新闻客户端成为媒体融合领域新的竞争高地，也成为媒体融合创新的重要赛道。在人工智能、NLP 自然语言处理、大数据和 5G 迅猛发展和全面应用落地的背景下如何赋能一款短视频客户端，尤其是围绕内容管理、音视频制作、审核发布、跨媒呈现等业务场景，如何创造性地去构建短视频新闻客户端的智能视频能力成为重要课题，需要我们为之探索实践。

2. 构建视频智能生产能力

目前，主流短视频新闻客户端存在 PGC 和 UGC 两种生产模式，因此两种不同的内容生产场景所需支撑的生产能力同样存在差异。从研发角度出发，首先要考虑内容生产能力的适用性，其次要考虑两种内容生产场景能力复用性，避免重复开发，从而降低研发成本。

结合人工智能、大数据、5G、VR、AR、MR 等技术赋能短视频新闻客户端形成“AI+”技术生态，主要有以下几个方面来构建智能音视频能力：

2.1　智能融合媒资系统

在融媒要求下，对于富媒体资源的使用场景存在跨平台、跨媒体、跨终端的刚性需求，需要利用云计算和人工智能，打造富媒体内容汇聚、存储管理、处理加工、渠道分发、分析挖掘为一体的云端智能媒资系统，它主要包括以下几方面的能力：

2.1.1　智能媒资存储

基于云计算弹性扩容能力、强大的中间件及丰富的数据库服务，为融合媒资搭建提供了底层能力保障，能支持各种媒体资料类型的存储，支持最多数据来源，可分场景管理，采用不同存储类型来支持不同业务需要。例如制作分发内容媒资，采用对象存储中的标准存储，它提供高可靠、高可用、高性能的对象存储服务，能够支持频繁的数据访问，适用于各种社交、分享类的图片、音视频应用等业务场景，另外同时也提供低频访问、归档存储、冷归档存储等多个存储类型，既能保证移动互联网高频、高并发访问的业务场景，同时又降低了使用成本。

2.1.2　智能媒资处理

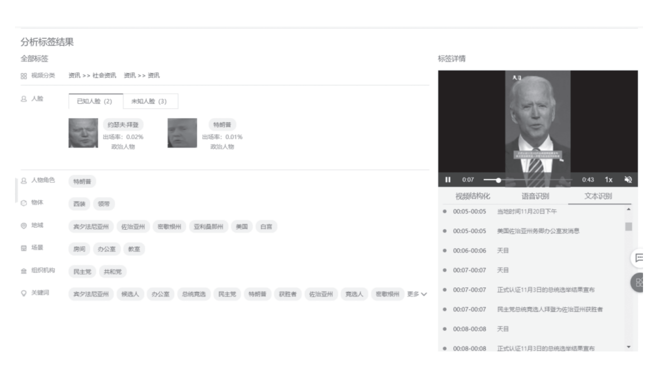
基于云存储之上，将音视频内容在云端进行一系列的多媒体数据处理，使音视频内容转码成适合在全平台播放的格式。须具备互联网音视频通用编码能力，如 H.264、H.265 编码，支持在输出视频上覆盖图片、文字水印，增强产品识别度。可将视频文件中单独分离出音

频或视频，对于长视频，具备视频分片并行转码，可大幅提升转码速度。对存储于对象存储上的视频文件可截取指定时间的 JPG 格式图像，支持单张截图、多张截图。可通过对视频内容的理解并结合画面美学，选出最优的关键帧作为视频封面图，或从视频内容中智能提取最有代表性的一组截图组成 GIF，作为视频的摘要概括。提取视频中的声音、图像及时序特征，生成视频指纹，实现视频片段溯源等功能，可用于视频去重、侵权视频过滤、原创视频保护等场景。

2.1.3　智能结构化管理

音视频内容是非结构化文件，在管理、使用、再次利用时存在查询、分析、挖掘等问题，传统花费大量的人力做编目管理，且编目管理存在多层编目要求，由于需要人工操作，除了费时费力以外，编目结果还存在操作人员对内容理解的差异性。

基于深度学习、计算机视觉技术和海量数据，对视频进行内容、文字、语音、场景等多模态分析，首先，可以实现自动输出新闻“5W”要素标签、视频分类和其他多维度的内容标签，适用于个性化推荐和视频搜索等场景；其次，识别视频中的人脸，并支持五官关键点定位、人脸属性分析和快速的人脸聚类，适用于机器剪辑、脱敏风控、人物关联、知识图谱等场景；再次，将音频转成文字，从文字中抽取关键词，也是标签内容的来源和脱敏风控的维度之一；最后，OCR 识别对多媒体数据中的文字进行检测、精准识别视频画面中的字幕、标题、弹幕等关键内容。



智能媒资结构化系统（媒立方 2.0 系统）

2.1.4　智能风控管理

在融媒体时代下大量的富媒体音视频内容的传播，传统的文字审核体系已经无法满足移动互联网时代海量内容的审核需求。通过人工智能机器审核 + 人工审核的方式，可有效降低内容审核人力成本，为音频、图像、视频内容安全提供了更高的技术保障，同时又提升内容审核的效率和准确度。该项能力主要提供涉政、涉黄、涉恐、涉暴、以及人物、场景、物体等多模态的内容风控能力。尤其是涉及 UGC 内容生产，内容数据多，增长快，且人工审核速度慢，成本高，存在巨大的违规风险，内容安全通过人工智能深度学习算法，自动化智能识别违规内容，大幅度降低人工成本，提高审核效率，能有效的满足风控管理要求。

2.1.5　PCDN 加速

以 P2P 技术为基础，通过挖掘利用边缘网络海量碎片化闲置资源构建低成本高品质内容分发网络服务，适用视频点播、直播等业务场景。通过集成 PCDN SDK 后，与普通 CDN 相比能显著提升分发质量，为实现视频“秒开”效果提供基础网络保证，一定程度也能降低分发成本。在安全防护机制上，对于防盗链、抵御 DDoS 攻击、数据安全方面使用加密鉴权制，节点缓存使用高强度加密，防止内容被篡改，确保内容受控。[2]

2.2　智能短视频制作平台

利用人工智能和计算机视觉技术，集素材管理、在线剪辑、后期包装、渲染导出和发布于一体的云在线制作平台，系统将视频素材在云端进行结构化处理，把视频中的元数据抽取出来进行自动化标签，即对上传的视频进行场景分类、人物识别、语音识别、文字识别等分析，形成层次化的精细分类标签，从而实现视频素材的精准搜索。通过智能特效、智能字幕、智能语音等 AI 技术提升在线视频制作效率，甚至基于 NLP 将脚本文字通过预设的制作模板快速转化为短视频。较常用的云快编功能可以让编辑直接基于浏览器快速剪辑视频添加转场特效，通过 AI 语音转文字快速添加字幕，利用 CV 技术实现虚拟主持人配音，渲染导出一键发布到客户端及各视频平台上。同时，平台还能将 5G 直播流实时接入，通过可视化操作，将直播视频以边播边打点的方式切片形成短视频立即输出发布。综上，编辑团队在线协同的生产方式，使整个视频制作的流程被大幅缩短，生产视频内容的门槛和成本也大大降低。AI 赋能视频制作，实现了视频处理生产力上的突破，改变了传统意义上视频制作的生产关系。



短视频智能快编系统（媒立方 2.0）

3. 5G 移动云直播系统

在 5G 网络大规模商用的情况下，基于 5G 高速率、低延时、万物互联的特性，为移动直播进入高清晰、强互动、全场景提供基础网络保证。首先，5G 高速率情况下可实现视频高码流，实现以高清画质为基础的高质量

信号传输，从而满足用户对融媒体直播越来越高的视觉体验。其次，5G 低时延特性使得新闻现场和用户感受之间实现“同步发生”，增强用户与媒体、用户与用户之间互动的欲望。再次，5G 让万物互联，从摄像机、无人机、360 VR 云台、交通安防监控到任意一台手机都可作为稳定的直播信号源，从而实现全场景、沉浸式直播内容。

综上所述，整套 5G 移动云直播系统主要包含以下几个方面：

1. 直播项目管理：支持直播项目创建、推流地址分配、直播时间设置、直播封面制作等功能。

2. 直播素材管理：对直播过程中使用到的各类素材进行集中管理，实现快速搜索、快速调用。

3. 实时导播管理：对多路直播信号流进行导播切换，实现跨时空、多机位、多场景画面管理。

4. 回放文件管理：对直播回放文件进行查看、下载、入库、回放替换、回放选段、回放快编。

5. 实时风控管理：实现对直播内容的播出控制，同时减少人为内容鉴定，提高内容智能化识别鉴定和预警。

6. 数据统计管理：对采集到的数据进行分析，实现对用户行为多维度分析并输出分析报表。



移动直播系统管理后台（媒立方 2.0）

4. 移动端视频应用能力

对于短视频客户端用户而言，最直观的就是视频播放体验，不管点播还是直播场景，视频播放器的加载方式、操作交互、视频清晰度、播放稳定性、以及流媒体文件解码的兼容性都关系到用户体验好坏与否。其次，具备 UGC 场景的短视频客户端须提供短视频拍摄、制作和特效功能，提升用户短频制作的趣味性，增强用户社交体验。

4.1 视频播放器

移动端播放器须支持下列技术要求：

1. 支持主流 MP4、M3U8、FLV、MKV 视频格式、MP3 音频格式以及 H.264、H.265 视频硬解码功能、AAC 音频编码，iOS 端须支持 AC3 音频编码，提供多码率 HLS 的无缝切换。

2.支持开始、结束、暂停、恢复、重播和循环播放等播放控制功能，同时提供点播和直播功能，支持网络视频的 URL 方式播放。

3.支持点播和转码的多路清晰度流切换，提供直播时移视频流播放和视频边播边缓存功能，适合短视频的循环播放场景。

4. 支持点播和直播的首屏秒开功能，提供直播的动态追帧，降低延时。具备直播的自动重连功能，已经缓冲的视频内容在拖动时不清除缓冲内容并快速拖动。

5.支持基于播放器用户行为信息的个性化埋点采集。

6. 支持实时指标监控、根因分析，能够提供全链路问题追踪能力。

4.2 短视频拍摄

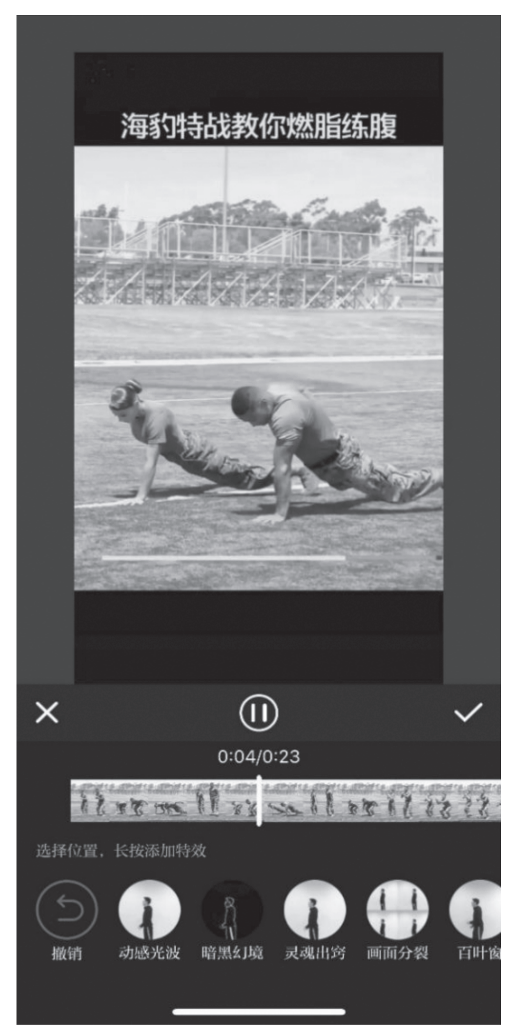
短视频拍摄功能须具备下列基础功能：

1. 支持断点录制、回删、点击拍摄、长按拍摄、闪光灯、实时水印、摄像头切换、分辨率设定、实时混音和变速等功能。

2.支持从相册选择视频、按视频时长和画面进行裁剪。支持多视频、多照片、照片和视频混合导入拼接、并可设置转场模式和转场持续时间。

3. 支持拍摄实时滤镜、普通美颜、人脸 2D 和 3D 贴图，提供基于人脸、人体的 CV 算法实现高级美颜、美妆、微整形、美体、手势识别等增强现实能力。

4.支持在编辑界面添加滤镜、字幕、动（静）态贴图、BGM 和涂鸦等功能。



天目新闻客户端 UGC 短视频制作工具

（下转第110页）